



NICK THIEBERGER
DIRECTOR OF PARADISEC
SCHOOL OF LANGUAGES AND LINGUISTICS
THE UNIVERSITY OF MELBOURNE
ARC CENTRE OF EXCELLENCE FOR THE DYNAMICS OF LANGUAGE

THE ONGOING CHALLENGE OF CONNECTING SPEAKERS TO ARCHIVAL LANGUAGE RECORDS

ICLDC 2019, UH MANOA

WHAT IS IN A LANGUAGE ARCHIVE AND HOW TO FIND IT

- ▶ Scale of documentation currently held and planned for the future
- ▶ Complicated role of archives (colonial or post-colonial?)
- ▶ Publicising items in collections to maximise chances of speakers finding it
- ▶ Building local collections

SCALING UP THE EFFORT

- ▶ “it is becoming feasible to record around 500 hours of linguistic material in the course of a year or two’s fieldwork, thanks to the miniaturisation, fidelity and portability of our recording devices.” (Evans 2017: 41)
- ▶ This will require
 - ▶ a change in fieldwork methods
 - ▶ more archives and increased capacity and efficiency of ingestion and delivery
 - ▶ 500 hours of audio and video is a terabyte or more of storage

LANGUAGE ARCHIVES

- ▶ Digital Endangered Languages and Musics Archives Network (DELAMAN)

- ▶ ALLA (Archive of the Indigenous Languages of Latin America)
- ▶ Alaska Native Language Archive
- ▶ California Language Archive
- ▶ ELAR (Endangered Languages Archive at SOAS)
- ▶ Kaipuleohone (University of Hawai'i Digital Ethnographic Archive)
- ▶ Native American Languages Collection at the Sam Noble Museum of Natural History



- ▶ PARADISEC (Pacific and Regional Archive for Digital Sources in Endangered Cultures)
- ▶ RWAAI (Lund University)
- ▶ The Language Archive (Max Planck Institute for Psycholinguistics)
- ▶ The Library of the American Philosophical Society (APS)
- ▶ World Oral Literature Project

LANGUAGE ARCHIVES

- ▶ Open Language Archives Community
 - ▶ Academia Sinica Collections
 - ▶ Alaska Native Language Archive
 - ▶ Archive of the Indigenous Languages of Latin America (AILLA)
 - ▶ C'ek'aedi Hwnax Ahtna Regional Linguistic and Ethnographic Archive
 - ▶ California Language Archive
 - ▶ CHILDES Data repository
 - ▶ COllections de COrpus Oraux Numeriques (CoCoON ex-CRDO)
 - ▶ Comparative Corpus of Spoken Portuguese
 - ▶ Endangered Languages Archive
- ▶ Kaipuleohone
- ▶ Living Archive of Aboriginal Languages
- ▶ PARADISEC
- ▶ Pacific Collection at the University of Hawai'i at Mānoa Hamilton Library
- ▶ PANGLOSS
- ▶ TALKBANK Data repository
- ▶ The Crúbadán Project
- ▶ The Language Archive at the MPI for Psycholinguistics
- ▶ The LDC Corpus Catalog
- ▶ Tibetan and Himalayan Digital Library

LANGUAGE ARCHIVES


- ▶ Organise linguistic records
- ▶ Depositing in an archive may be the first time that a linguist organises their collection
- ▶ Provide a citable form of language records
- ▶ Curate records with licenses for use
- ▶ Make records available to source communities
- ▶ Are always underresourced and so need to prioritise their activities

SAMPLE OF LANGUAGE ARCHIVE STATISTICS (OLAC)

Archive	Number of Resources	Distinct Languages
Archive of the Indigenous Languages of Latin America (AILLA) (USA)	289	257
C'ek'aedi Hwnax Ahtna Regional Linguistic and Ethnographic Archive (USA)	1,474	3
California Language Archive (USA)	14,582	281
COllections de COrpus Oraux Numeriques (CoCoON ex-CRDO) (France)	13,437	217
Endangered Languages Archive (SOAS, UK)	89,575	463
Kaipuleohone (USA)	4,785	199
The Language Archive (Nijmegen, Netherlands)	127,985	433
Living Archive of Aboriginal Languages (Australia)	2,966	38
Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) (Australia)	17,084	1202
Pacific Collection at the University of Hawai'i at Mānoa Hamilton Library (USA)	5,292	749
Total	277,469	

OPEN LANGUAGE ARCHIVES COMMUNITY (OLAC)

<http://www.language-archives.org/language/erk>
Using ISO-639-3 codes for languages



OLAC resources in and about the South Efaté language

ISO 639-3: [erk](#)

The combined catalog of all OLAC participants contains the following resources that are relevant to the language:

- [Primary texts](#)
- [Lexical resources](#)
- [Language descriptions](#)
- [Other resources about the language](#)
- [Other resources in the language](#)

Other known names and dialect names: South Efaté, Erakor, Fata, Southern Efaté

Use faceted search to [explore resources for South Efaté language](#).

Primary texts

1. [ONLINE](#) *Fila Story: A Oa Erua; B&C: Lamenu (cont.); Baki; Pango-Erakor-Eratap*. Arthur Capell (compiler); Arthur Capell (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC1-203](#)
2. [ONLINE](#) *Oriya, Lowolwol*. Arthur Capell (compiler); Arthur Capell (researcher). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC1-411](#)
3. [ONLINE](#) *Prodigal Son in Efaté Dialects*. Arthur Capell (compiler); Arthur Capell (researcher). 1900. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC1-416](#)
4. [ONLINE](#) *Texts in languages from Efaté*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT403](#)
5. [ONLINE](#) *The Prodigal Son story in Samoa Point/Havannah Harbour*. Arthur Capell (compiler); Arthur Capell (researcher). 1900. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT301](#)
6. [ONLINE](#) *The Gospel According to Luke in the language of Efaté*. Arthur Capell (compiler); Arthur Capell (researcher). 1900. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT501](#)
7. [ONLINE](#) *Catechism in the language of Efaté*. Arthur Capell (compiler); Arthur Capell (researcher). 1900. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT502](#)
8. [ONLINE](#) *Notes on Nguná-Efaté dialects*. J. Graham Miller. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT104](#)
9. [ONLINE](#) *South Efaté Language Primer*. Arthur Capell (compiler); Arthur Capell (researcher); Peter Newton (depositor). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT101](#)
10. [ONLINE](#) *Notebook containing information about Efaté and other South Vanuatu languages*. Arthur Capell (compiler); Arthur Capell (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT301](#)
11. [ONLINE](#) *Prodigal son text in Efaté*. Arthur Capell (compiler); Arthur Capell (researcher); Arthur Capell (depositor). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT306](#)
12. [ONLINE](#) *Prodigal Son text in Erakor*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT307](#)
13. [ONLINE](#) *Prodigal son text in Eratap*. Arthur Capell (compiler); Arthur Capell (researcher); Arthur Capell (depositor). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT308](#)
14. [ONLINE](#) *CD01*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT1-001](#)
15. [ONLINE](#) *Recordings in South Efaté*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT1-002](#)
16. [ONLINE](#) *CD03*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT1-003](#)
17. [ONLINE](#) *Recordings in South Efaté*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT1-004](#)

Lexical resources

1. [ONLINE](#) *Bible texts in several Vanuatu languages*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VCOMVQ002](#)
2. [ONLINE](#) *Efaté dialects*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT401](#)
3. [ONLINE](#) *Efatése Grammar and dictionary*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT402](#)
4. [ONLINE](#) *Spirax note book*. N.H. Ill. Arthur Capell (compiler); Arthur Capell (researcher); Peter Newton (depositor). 1900. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VFN15](#)
5. [ONLINE](#) *Notebook containing information about Efaté and other South Vanuatu languages*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT301](#)
6. [ONLINE](#) *Topics in the grammar and documentation of South Efaté, an Oceanic language of Central Vanuatu*. Nick Thieberger. 2005. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT2-PHD](#)
7. [ONLINE](#) *Sesake vocabulary recorded by Bishop Patteson, 1866*. Nick Thieberger (compiler). 1866. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT3-SESAKE](#)
8. [ONLINE](#) *Dictionary and texts in South Efaté*. Nick Thieberger (compiler); Nick Thieberger (depositor). 2006. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT8-TEXT](#)

Language descriptions

1. [ONLINE](#) *Glottolog 2.3 Resources for South Efaté*. n.a. 2014. Max Planck Institute for Evolutionary Anthropology. [oai:glottolog.org:sout2856](#)
2. [ONLINE](#) *Efatése Grammar and dictionary*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT402](#)
3. [ONLINE](#) *The Nguná-Efaté dialects*. Arthur Capell (compiler); Arthur Capell (researcher). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT11](#)
4. [ONLINE](#) *Notes on Tongos language*. Arthur Capell. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:AC2-VSEFAT207](#)
5. [ONLINE](#) *Topics in the grammar and documentation of South Efaté, an Oceanic language of Central Vanuatu*. Nick Thieberger. 2005. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT2-PHD](#)
6. [ONLINE](#) *PHOIBLE Online phonemic inventories for South Efaté*. n.a. 2014. Max Planck Institute for Evolutionary Anthropology. [oai:phoible.org:erk](#)
7. [ONLINE](#) *WALS Online Resources for Efaté (South)*. n.a. 2008. Max Planck Institute for Evolutionary Anthropology. [oai:wals.info:erk](#)

Other resources about the language

1. [ONLINE](#) *Topics in the grammar and documentation of South Efaté, an Oceanic language of Central Vanuatu*. Thieberger, Nicholas. 2004. WALS Online RefDB. [oai:wals.info:erk](#)
2. [ONLINE](#) *Houses in Erakor*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-02](#)
3. [ONLINE](#) *Going away party*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-03](#)
4. [ONLINE](#) *Going away party*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-04](#)
5. [ONLINE](#) *Going away party*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-05](#)
6. [ONLINE](#) *Kalsarap holding a rake (nkur) made from the tree of the same name*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer); Kalsarap Namaf (participant). 2001. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-06](#)
7. [ONLINE](#) *Yokopeth*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer); Yokopeth (participant). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-1_03](#)
8. [ONLINE](#) *Apu and Limas reading book and dictionary*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer); Kalsarap Namaf (participant); Limas (participant). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-1_04](#)
9. [ONLINE](#) *Nick Wayne being held by an unknown young woman*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer); Nick Wayne (participant). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-1_05](#)
10. [ONLINE](#) *Metu Josef*. Nick Thieberger (compiler); Nick Thieberger (depositor); Nick Thieberger (photographer). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au:NT4-1_06](#)

PACIFIC AND REGIONAL ARCHIVE FOR DIGITAL SOURCES IN ENDANGERED CULTURES (PARADISEC)

- ▶ Collaborative digital research resource set up by University of Sydney, University of Melbourne & Australian National University, 2003 with the aim of preserving ethnographic (language, oral history, music) recordings made since the 1950s



funding from Australian Research Council
LIEF Scheme (2003-5), Grangenet (2005-6), APSR
(2006-7), LIEF (2011), ARC Centre of Excellence for the
Dynamics of Language (2014-2021)




PARADISEC

- ▶ 1,203 languages
 - ▶ 489 collections
 - ▶ 22,937 items
 - ▶ 273,160 files
 - ▶ 9,650 hours (audio)
 - ▶ 820 hours (video)
 - ▶ 55 TB of files
-
- ▶ 30.7 TB of video files in archive
 - ▶ 14.8 TB of audio files in archive
 - ▶ 1.29 TB of image files in archive
 - ▶ 1.16 GB of text files in archive
-
- ▶ January 2019



API, METADATA FEEDS



OLAC: Open Language Archives Community

[HOME](#) | [DOCUMENTS](#) | [ABOUT](#) | [ARCHIVES](#)
[NEWS](#) | [ORGANIZATION](#) | [TOOLS](#) | [SERVICES](#)


SEARCH THIS SITE:

OLAC Mission

OLAC, the Open Language Archives Community, is an international partnership of institutions and individuals who are creating a worldwide virtual library of language resources by: (i) developing consensus on best current practice for the digital archiving of language resources, and (ii) developing a network of interoperating repositories and services for housing and accessing such resources.

Find Language Resources

Try OLAC's new search engine at: <http://search.language-archives.org/>



Archive: -- all archives --

Region: [Africa](#) [Americas](#) [Asia](#) [Europe](#) [Pacific](#)

News

OLAC Joins the Linguistic Linked Open Data Cloud: The OLAC system has now been integrated with LLOD... [More...](#)

New OLAC Page Listing Archive Submission Policies: As a service to linguists who are in search of an archive that could receive a deposit... [More...](#)

New OLAC Search Service: In December 2010, to mark our 10th anniversary, OLAC announces a new search service... [More...](#)

Language Archives Tutorial at LSA: In January 2009, an OLAC-endorsed tutorial on language archives was held at the winter meeting of the Linguistic Society of America... [More...](#)

OLAC Presented at three Pacific Conferences: In 2008 and 2009, OLAC is being presented at conferences in Australia, the Philippines, and Hawaii... [More...](#)

[More news...](#)

Documents

[OLAC Standards](#) - specify how OLAC operates

[Recommendations](#) - express consensus of OLAC members regarding language resource archiving

[Notes](#) - background information and guidance for implementers

General Information:
[Overview](#) | [FAQ](#) | [Implementers' FAQ](#)

Participating Archives • OLAC • Delivered by the Penn Libraries
Printer-Friendly Page

OLAC Language Resource Catalog

Search for language resources

Results:
Showing hits 1 - 50 out of 11142

Currently Used Filters
✓ Archive: Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC)

Sort Results By:
Possible Sorts:
• Title [a-z][z-a]
• Id [a-z][z-a]
• Date [a-z][z-a]

Narrow Results By:
Online browse
• Yes 11055
• No 87
Subject language browse
• No linguistic content 649
• South West Bay 403
• Bahinemo 400
• South Efate 396
view more...
Language family browse
• Austronesian 4106
• Malayo-Polynesian 3883
• Eastern Malayo-Polynesian 3327
• Papuan 2839
view more...
Geographic region browse
• Pacific 7075
• Asia 2000
• Europe 419
• Africa 271
view more...
Country browse
• Papua New Guinea 3543
• Vanuatu 1740
• Indonesia 1349
• Australia 1087

Navigating the Catalog

- Catalog Home
- Search Strategies
- Advanced Search
- New: Records recently added or modified

Quick Links

- Browse by Language
- Browse by Country
- Browse by Linguistic Field
- Browse by Linguistic Type
- Browse by Language Family

Contacts

- Email Us

More information

- OLAC Homepage
- OLAC FAQ
- Participating Archives

Powered by the DLA

Taurani Rua Araara 6 2012
Sally Nicholas (compiler); Sally Nicholas (speaker); Tuakana Taurani (speaker); Taurereau Taurani (speaker). 2012. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

How the birds were given their responsibilities
Cindy Schneider (compiler); Cindy Schneider (depositor); Joseph Bulemeres (speaker). 2003. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

27 Fale at Falealupo
David Goldsworthy (compiler); David Goldsworthy (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

Nadrogā
Albert Schütz (compiler); Albert Schütz (depositor); Albert Schütz (recorder). 1961. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

Nafasiri; Wainimata; Naivucini; Naitavoli
Albert Schütz (compiler); Albert Schütz (recorder). 1961. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

Eastern Arrernte vocabulary and discussion 2
Gavan Breen (compiler). 1986. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

Two Namakura stories
David Luders (compiler); Pakoasongi Parang (speaker). 1996. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

Hunting8
Wayne Dye (compiler); Wayne Dye (recorder). 2008. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

nagay imbuu
Laura Dimock (compiler); Laura Dimock (researcher). 2007. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

4A
Don Laycock (compiler); Don Laycock (recorder). 1967. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).

Feeds picked up by ANDS, OAI services, Open Language Archives Community (OLAC), etc.

<http://search.language-archives.org/search.html?>

[fq=archive_facet%3A%22Pacific%20And%20Regional%20Archive%20for%20Digital%20Sources%20in%20Endangered%20Cultures%20%28PARADISEC%29%22](http://search.language-archives.org/search.html?fq=archive_facet%3A%22Pacific%20And%20Regional%20Archive%20for%20Digital%20Sources%20in%20Endangered%20Cultures%20%28PARADISEC%29%22)

HOW TO GET RECORDS BACK TO THEIR SOURCE?

- sending hard disks to cultural centres
- usb drives to plug into xbox or tv (perhaps not for computers)
- providing low resolution copies of archival files for delivery
- making our catalog maximally accessible and findable
- problem of providing context for a set of files
- google searching is the most likely way that speakers will find our records

ARCHIVES CRITIQUED AS INACCESSIBLE TO SPEAKERS

- ▶ “One major insight from the workshop was the discovery that at present, most language archives are not meeting the needs of most users.” Wasson et al 2018: 642
- ▶ We can all do better, however, PARADISEC has many users and they are spread through remote and urban areas

ARCHIVES CRITIQUED AS INACCESSIBLE TO SPEAKERS

- ▶ “Archives have made explicit attempts to serve this community; however, the vast majority of these attempts have been top-down in nature, inspired by the archives’ vision of the needs of these user communities.” (Dobrin & Holton 2013)
- ▶ Another way of seeing this is that linguists are taking responsibility for repatriation
- ▶ Would a preferable response be the one that archives are always criticised for, locking up the data in inaccessible forms?

-
- ▶ “The challenge is to design language archives that accommodate the needs of all user groups. A user-centered design process could help make language archives more accessible and more useful to larger numbers of users.” (Wasson et al 2016: 643)
 - ▶ Meantime, archives need to keep accessioning and making material available

USERS FINDING THEIR LANGUAGE RECORDS

Searching for his language 'Koitabu': second page of a google search

Motu-Koitabu electoral roll - The National

<https://www.thenational.com.pg/motu-koitabu-electoral-roll/> ▼

Apr 26, 2018 - I make reference to a report in this newspaper on the Motu-**Koitabu** electoral roll and the removal of 3000 foreign or non-native names from the ...

Motu Koitabu Assembly Act Never Recognised says Acting Chairman ...

<https://emtv.com.pg/motu-koitabu-assembly-act-never-recognised-says-acting-chairm...> ▼

The Motu **Koitabu** Assembly Act 2007 has never been recognized by the Government since its...

Motu Koitabu Environment and Health ... - Project Detail - UNDP

<https://sgp.undp.org> › Projects › Project Search Results ▼

A planning grant of K1,966.67 was provided to the Motu **Koitabu** Group of the Village of Hanuabada to develop their project proposal. In late 1998 the group ...

OLAC resources in and about the Koitabu language

www.language-archives.org/language/kqi ▼

Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).
oai:paradisec.org.au:SAW1-018; ONLINEMotu, **Koitabu**: LMS Singing ...

USERS FINDING THEIR LANGUAGE RECORDS

If only you witnessed and captured the reaction in me going through the recordings at home! It is quite an amazing experience! From feeling of awe to emotion to deep excitement!

The feeling of knowing that your language has been documented or recorded in a structured way, kept safely somewhere in the world, hearing it spoken 50 - 60 years ago and by some people you haven't seen but whose names you only hear in history etc is quite incredible.

It is most heartwarming to know that it is possible to sustain the life of my language. Thank you once again for the opportunity to listen to the records!

As I see it now, it is "US OR NEVER!" In about three or four Koitabu villages, the observation is that people in my peer group can understand in full but speak a little. Generation after us cannot do both. I believe that if we don't teach ourselves now and pass on, the language will be totally gone in the next 10 to 20 years.

E'ava Geita, Koitabu, PNG, 23/3/2015

ENGLISH ONLY?

- ▶ Of all the language archives listed earlier, most present English only interfaces to the collection, and English only metadata
- ▶ The Archive of the Indigenous languages of Latin America, AILLA, provides a Spanish interface (and metadata labels)
- ▶ COllections de COrpus Oraux Numériques, CoCoON, is French only
- ▶ Pangloss is French and English
- ▶ PARADISEC is English only, but allows metadata in any language and has information brochures in French, Mandarin, Indonesian, Japanese, Tok Pisin and Bislama

LOW RESOLUTION COPIES

- ▶ Archival files in PARADISEC are transcoded to compressed versions on ingest
- ▶ wav > mp3
- ▶ tif > jpg
- ▶ video > mp4
- ▶ Both versions are provided in the collection

Folk Stories



Multimedia item 16 of 18



DOWNLOAD MP3



0:09 / 2:16



EAF

▶ PLAY (0:02)



A_Transcription-txt-mos

are? n̄ cinlu c^ha bə vu c^ha bə aran c^ha ate c^ha. po? bə
mun̄kaha? n̄ ŋa ba doŋ k̄ŋ ŋamə twez̄u pan^hum.

Cinlu is a small bird, the elephant thinks, 'I am largest
animal in the world.'

▶ PLAY (0:15)



A_Transcription-txt-mos

<https://github.com/marcolarosa/pdsc-collection-viewer>

PUBLICISING THE COLLECTION

- ▶ Publicity helps lead both to discovery of what is in the collection, and to accession of new material
- ▶ PARADISEC Facebook page has 1,933 followers

PUBLICISING THE COLLECTION



Glossopticon : <https://www.youtube.com/watch?v=X2iimqq7LS8>

<https://glossopticon.com/vr.html>

MAKING SUBCOLLECTIONS FOR DELIVERY

**PARADISEC
collection**

20,911 items

238,977 files

Delivery to source
community

Item 1 (20 files)

Item 2 (2 files)

Item 3 (4 files)

Item 4 (7 files)

Problem : no catalog
accompanies the files

FILE DUMP

NT10-W11-A.wav.txt

NT10-W11-A.txt

NT10-W09-A.txt

NT10-W08-BAMP.txt

NT7-200805-1.mp3

NT7-200804-1.mp3

NT7-200803-1.mp3

NT7-200801-4.mp3

NT7-200801-1.mp3

NT6-200518.mp3

NT6-200517-A

NT6-200515-A.mp3

NT6-200513.mp3

NT6-200513-A.mp3

NT6-200511-A.mp3

NT6-200511-A

NT5-TokelauOf-vid.mov

NT5-Tokelau-Vid1.mov

NT5-StringBand-007.dv

NT5-StringBand-001.mov

NT5-DickLauto-Vid2.mov

NT5-200801-4.wav

NT5-200801-3.wav

NT5-200605-A.wav

NT5-200604.wav

NT5-200602.wav

NT5-200601.wav

NT5-200519-A.wav

NT5-200518-A.wav

NT5-200513-A.wav

NT5-200501.wav

NT1-98021-A

NT1-98020-B

PARADISEC CATALOG WRITES XML FILES

```
<?xml version='1.0' encoding='utf-8' ?>
```

```
<item>
```

```
<identifier>NT1-006</identifier>
<collectionId>154</collectionId>
<private>>false</private>
<title>Recordings in South Efate</title>
<description>Stories of Langtatalof, and Katap
<originationDate>1997-11-11</originationDate>
<originationNarrative>1997-10-23, 1997-11-11</
<archiveLink>http://catalog.paradisec.org.au/c
<url/>
<collector>Nick Thieberger</collector>
<countries>
  <country>Vanuatu - VU</country>
</countries>
<language>Nafsan</language>
<subjectLanguages>
  <language>Efate, South - erk</language>
</subjectLanguages>
<contentLanguages>
  <language>Bislama - bis</
  <language>Efate, South -
</contentLanguages>
<dialect/>
<region>Erakor village</regio
<geographicLocation>
  <south_limit>-17.8235</so
  <north_limit>-17.7235</no
  <east_limit>168.317</east
  <west_limit>168.217</west
</geographicLocation>
<university>University of Melbourne</universit
<operator>Corin Bone</operator>
<dataCategories>
  <category>primary text</category>
  <category>sound</category>
</dataCategories>
<discourseType/>
<agents>
```

```
<archiveInfo>
  <bornDigital>>false</bornDigital>
  <tapesReturned>>false</tapesReturned>
  <originalMedia>audiocassette</originalMedia>
  <dateReceived/>
  <dateDigitised>2001-01-01 00:00:00 +1100</dateDigit
  <ingestNotes/>
  <metadataImportedOn/>
  <metadataExportedOn/>
  <tracking>1 CD given to the Vanuatu Cultural Centre
</archiveInfo>
<adminInfo>
  <recordCreated>2012-09-27 20:08:00 +1000</recordCre
  <recordLastModified>2016-03-29 02:13:43 +1100</reco
  <adminAccess> </adminAccess>
  <dataAccessConditions>Open (subject to agreeing to
```

```
<agent role="depositor">
  Nick Thieberger
</agent>
  <agent role="recorder">
    Nick Thieberger
  </agent>
  <agent role="speaker">
    Kalsarap Namaf
  </agent>
  <agent role="speaker">
    Iokopeth
  </agent>
</agents>
<citation>Nick Thieberger (collector), Nick Thieberger (depositor,
```



NT1-006-006A.mp3



NT1-006-006A.wav



NT1-006-CAT-PDSC_ADMIN.xml

```
<collector>Nick Thieberger</collector>
<operator/>
<university>University of Melbourne</university>
<countries>
  <country>Vanuatu - VU</country>
</countries>
<languages>
  <language>Bislama - bis</language>
  <language>Efate, South - erk</language>
  <language>Lelepa - lpa</language>
</languages>
```

```
</file>
<file>
  <name>NT1-006-006A.wav</name>
  <mimeType>audio/x-wav</mimeType>
  <fileSize>331 MB</fileSize>
  <duration>00:32:47.569</duration>
  <bitrate>1.35 MB</bitrate>
  <sampleRate>44100 Hz</sampleRate>
  <channels>Stereo</channels>
  <frameRate/>
</file>
</files>
```

DOWNLOADER

- ▶ Harvests individual catalog xml files and creates a static html page catalog of just that assemblage of items
- ▶ We had previously built a javascript viewer for items in PARADISEC and re-used that to be the browser
- ▶ After testing, we added new functionality to group items by genre
- ▶ <http://www.language-archives.services/about/data-loader>

ADDING GENRE AS A METADATA TERM

- ▶ Adapting the catalog to suit a particular requirement
- ▶ Comments field used to group items by genre
- ▶ Captured by the dataloader for delivery on raspberry pi

Admin Information

Record created 2018-08-07 16:26:12 +1000

Record modified 2018-08-09 02:01:08 +1000

Comments [genre: MIRLPA :::]

CAT-PDSC file [View item XML file](#)

Admin Information

Record created 2018-08-10 14:19:32 +1000

Record modified 2018-08-10 14:40:41 +1000

Comments [genre: TJILKUKU :::]

CAT-PDSC file [View item XML file](#)

RASPBERRY PI



- ◆ Delivery of archival files avoiding bandwidth cost
- ◆ Raspberry Pis are low powered webserver
 - wifi transmitter
 - USB drive stores the data

<http://www.language-archives.services/>

<https://github.com/marcolarosa/pdsc-librarybox-installer>

- ▶ Mobile phones – 80% coverage in most Melanesian countries



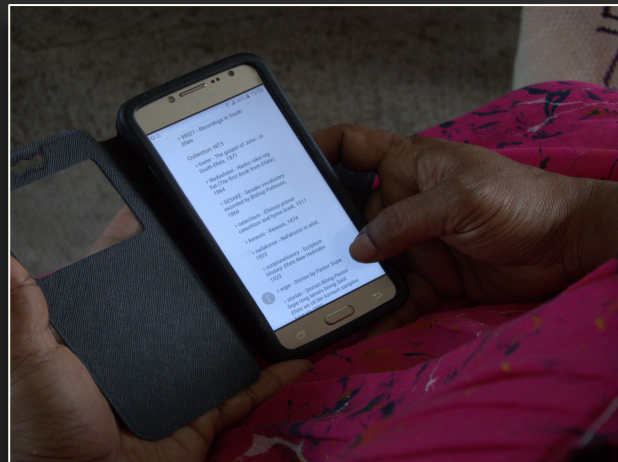
Port Vila, Vanuatu, 'Internet on your phone', 2012

ACCESS VIA MOBILE PHONES

- ▶ Tested in Vanuatu, Tahiti, Western Desert (Australia)



Elisa Alphonse in Erakor village



Kalonuk Albert & Gino Joseph in Erakor village

ARCHIVES PROBABLY REQUIRE LOCAL INTERMEDIARIES

- ▶ Archives should have the preservation of primary records as their main goal.
- ▶ If someone can navigate google then they are probably going to find records in their language and can then negotiate the interface of some, if not all, current language archives
- ▶ For those who are unable to search the web, local agencies can present archival information in more relevant ways
 - ▶ language centres, cultural centres
- ▶ These are where the raspberry pi can sit and broadcast for phone users

REFERENCES

- ▶ Dobrin, Lise & Gary Holton. 2013. The documentation lives a life of its own: The temporal transformation of two endangered language archive projects. *Museum Anthropology Review* 7. 140-154.
- ▶ Evans, Nicholas. 2017. Ngûrrahmalkwonawoniyan: Listening Here. *Humanities Australia: The Journal of the Australian Academy of the Humanities* 8 (2017): 34-44.
- ▶ Wasson, Christina, Gary Holton & Heather S. Roth. 2016. Bringing User-Centered Design to the Field of Language Archives. *Language Documentation & Conservation* 10. 641-681.



Australian Research Council – LIEF
grants LE110100142, LE0560711,
LE0453247

ARC DP0450342, DP0984419,
FT140100214, CE140100041



ARC CENTRE OF EXCELLENCE FOR
THE DYNAMICS OF LANGUAGE